

大肠杆菌 *yigP* 基因转录调控序列的鉴定

汪屹, 叶江, 张惠展*

华东理工大学生物反应器工程国家重点实验室, 上海 200237

摘要: 【目的】调查 *yigP* 基因启动子的活性, 并对该转录调控序列进行分析。【方法】以 *lacZ* 为报告基因, 克隆启动子片段至启动子探针质粒中, 通过检测 β -半乳糖苷酶活性判断启动子活性, 并通过克隆一系列逐步缩短的启动子片段来确定启动子所在区域。利用定点突变技术, 对启动子的重要序列进行定点突变, 调查其对启动子活性的影响。【结果】确定了 *yigP* 基因启动子的区域, 鉴定了启动子的-10区和-35区, 并发现了启动子上游存在一个负调控序列, 对该序列进行了初步的研究显示其中部分序列是这种负调控作用的核心序列。【结论】对 *yigP* 基因的转录调控序列进行了鉴定, 丰富了我们对基因转录调控的认识。

关键词: *yigP*, 转录调控, 启动子, 报告基因, 点突变

中图分类号: Q933 文献标识码: A 文章编号: 0001-6209 (2012)05-0566-07

基因表达调控是现代分子生物学的中心课题之一。基因表达调控主要表现在以下几个方面: 第一, 转录水平上的调控; 第二, mRNA 加工、成熟水平上的调控; 第三, 翻译水平上的调控。在转录水平上的调控简称为转录调控, 它涉及遗传信息从基因组 DNA 向 RNA 的传递过程, 包括染色体上的顺式作用元件与反式作用因子之间的相互作用, 从而介导基因转录的时空特异性和转录水平。对于顺式作用元件的研究有助于理解基因的转录调控。

启动子是 DNA 分子上能与 RNA 聚合酶结合并形成转录起始复合体的区域, 它能够控制基因的转录。在原核生物中, σ^{70} 识别的启动子包含两个短序列, 即位于从转录起始点起计的-10及-35上游位置, 分别称为-10区和-35区, 其序列相对保守, 其中前者的保守序列为 TATAAT, 后者的保守序列为 TTGACA^[1]。在大肠杆菌中, σ^{70} 依赖的启动子除了

有上述保守的-10区以外, 通常也会含有一个名叫“extended -10”的元件^[2], 这种元件起先是在枯草杆菌中被发现的。这种“extended -10”元件通常位于启动子-10区的上游相隔一个碱基处, 其主要的特征序列为 5'-TG-3', 该元件对启动子活性的影响很大, 任何一个碱基突变都将导致启动子活性的急剧下降。据文献报道, 这种启动子在大肠杆菌中约占 20%^[3]。在革兰氏阳性细菌中, 这种序列更为保守, 表现为 5'-TRTC-3'^[4]。

基因的表达调控不仅与蛋白质等反式作用因子相关, 而且也受 DNA 的顺式调控元件影响, 或激活转录, 或抑制转录^[5]。例如操纵基因, 它是与一个或者一组结构基因相邻近, 并且能够与一些特异的阻遏蛋白相互作用, 从而控制邻近的结构基因表达的基因。

李月^[6]在对 *yigP* 基因的研究过程中发现该基

基金项目: 国家自然科学基金 (31070073)

* 通信作者。Tel: +86-21-64252515; E-mail: huizhzh@ecust.edu.cn

作者简介: 汪屹 (1986-), 男, 浙江宁波人, 硕士研究生, 研究方向为分子生物学。E-mail: wangyi495@126.com

收稿日期: 2012-01-09; 修回日期: 2012-02-18

因是大肠杆菌生长必需基因,而且只需回补该基因的后半截片段 P4P2 就足以使细菌正常生长。并且我们后期的研究发现 *yigP* 基因的 P4P2 片段具有自身启动子活性。因此本论文利用以 *lacZ* 为报告基因的启动子探针质粒 pSP-Z^[7],对 *yigP* 基因的启动子序列进行了研究。大致定位了启动子的区域,并结合点突变技术,对 *yigP* 启动子的-10 区和-35 区进行了鉴定。还发现了启动子的上游含有一段负调控序列,对该负调控序列进行了初步调查。

1 材料和方法

1.1 材料

1.1.1 菌株:本文所用的菌种为大肠杆菌 JM83。

1.1.2 培养基和抗生素:大肠杆菌培养采用 Luria-Bertani 培养基,氨苄青霉素 (Ap) 使用终浓度为 100 $\mu\text{g}/\text{mL}$ 。

1.1.3 引物:本文所用引物及其特性见表 1。

表 1 本研究所用的引物
Table 1 Primers used in this study

Primers	Sequence (5'→3')	Remark
P40	GAAGCTTACCGCACTGATTCC	HindIII
P41	CCCAAGCTTCAGTGGTGAGCTGGAAG	HindIII
P42	CCCAAGCTTTGCAGGGCGATATTCAG	HindIII
P43	CCCAAGCTTGTGGTGCAAACTTCG	HindIII
P44	CCCAAGCTTCGTTGCGCTGGCAGATC	HindIII
V3	GATTCCTCCATGGCGGATATCACCG	NcoI
Pro-35	ATCTGCCAGCGCAACGAAGTTTTCACC	
Pro-101	TTGCGCTGGCAGATCTGGTATAATTCGACCCCTGCGGAA	mutational site
Pro-102	TTGCGCTGGCAGATCTGGCAGAGCTCGACCCCTGCGGAA	mutational site
LLV3	TAATTACCATGGGGGATATCACCGGTATAAGGGGCCAGCAGTCCCGCAGGGTCCA	NcoI
P436	AACCCCAAGCTTGTGGTGCAAAAC	HindIII
Pro-351	AGATCTGCCAGCGCTGTCAAGTTTTCACCACAA	mutational site
Pro-352	AGATCTGCCAGCGCAACGCCGTTTTCACCACAA	mutational site
Pro-10	GCGCTGGCAGATCTGGCAGAGTTTCGACCCCTGCGGAA	
Pro-1G1	TTGCGCTGGCAGATCACGCAGAGTTTCGACCCCTGCGGAA	mutational site
P42M01	GTACGTAAGCTTAGCAGGGCGATATTCAG	HindIII
P42M02	GTACGTAAGCTTCCAGGGCGATATTCAG	HindIII
P42M03	GTACGTAAGCTTTGGAGGGCGATATTCAG	HindIII
P42M04	GTACGTAAGCTTTGCTGGGGCGATATTCAG	HindIII
P42M05	GTACGTAAGCTTTGCAGGGCGATATTCAG	HindIII
P42M06	GTACGTAAGCTTTGCAGGGCGATATTCAG	HindIII
P42M07	GTACGTAAGCTTTGCAGGGCGATATTCAG	HindIII
P42M08	GTACGTAAGCTTTGCAGGGGGATATTCAGG	HindIII
P42M09	GTACGTAAGCTTTGCAGGGCCATATTCAGGT	HindIII
P42M10	GTACGTAAGCTTTGCAGGGCGTTATTCAGGT	HindIII
P42M11	GTACGTAAGCTTTGCAGGGCGAAATTCAGGTGG	HindIII
P42M12	GTACGTAAGCTTTGCAGGGCGATTTTCAGGTGGT	HindIII
P42M13	GTACGTAAGCTTTGCAGGGCGATAATTCAGGTGGTG	HindIII
P42M14	GTACGTAAGCTTTGCAGGGCGATATTCAGGTGGTG	HindIII
P42M15	GTACGTAAGCTTTGCAGGGCGATATTCAGGTGGTGCA	HindIII
P42M16	GTACGTAAGCTTTGCAGGGCGATATTCAGGTGGTGCAA	HindIII
P42M17	GTACGTAAGCTTTGCAGGGCGATATTCAGGTGGTGCAAA	HindIII

1.1.4 主要试剂和仪器:DNA 限制性内切酶、DNA 连接酶、高保真 DNA 聚合酶及 Taq DNA 聚合酶、DNA MARKER DL2000、pMD18-T 载体均购自 TaKaRa 公司;PCR 在 Eppendorf 公司(德国)生产的

Mastercycler personal PCR 仪上进行;凝胶成像系统由上海复日科技有限公司生产。

1.2 DNA 的提取和转化

大肠杆菌感受态细胞制备和质粒 DNA 提取、转

化等常规 DNA 操作参照分子克隆手册。

1.3 质粒的构建

1.3.1 启动子上游边界确定质粒的构建:以大肠杆菌 JM83 基因组 DNA 为模板 PCR 扩增一系列不同长度的片段,上游引物分别为 P40, P41, P42, P43, P44, 下游引物均为 V3。将 PCR 获得的片段,克隆到 pMD18-T 上,再经 *Hind*III 和 *Nco*I 联合酶切后克隆至经相同酶切的载体质粒 pSP-Z 上,分别得到质粒 pP40V3-Z, pP41V3-Z, pP42V3-Z, pP43V3-Z, pP44V3-Z。

1.3.2 含启动子-10 区, -35 区和 TG 突变质粒的构建:分别用如下引物组合 (P43, Pro-35, Pro-101/Pro-102, LLV3); (P436, Pro-351/Pro-352, Pro-10, LLV3); (P43, Pro-35, Pro-TG1, LLV3), 利用各组引物之间的互为模板进行 PCR 延伸、扩增, 所获得的片段, 经 *Hind*III 和 *Nco*I 联合酶切后克隆至经相同酶切的载体质粒 pSP-Z 上, 分别得到质粒 pP43V3-Z101, pP43V3-Z102, pP43V3-Z351, pP43V3-Z352, pP43V3-ZAC。

1.3.3 负调控序列区域点突变质粒的构建:以大肠杆菌 JM83 基因组 DNA 为模板 PCR 扩增一系列含不同突变的片段, 上游引物分别为 P42M01 → P42M17, 下游引物均为 V3。将 PCR 获得的片段, 经 *Hind*III 和 *Nco*I 联合酶切后克隆至经相同酶切的载体质粒 pSP-Z 上, 分别得到质粒 pP42V3-ZMut01 → pP42V3-ZMut17。

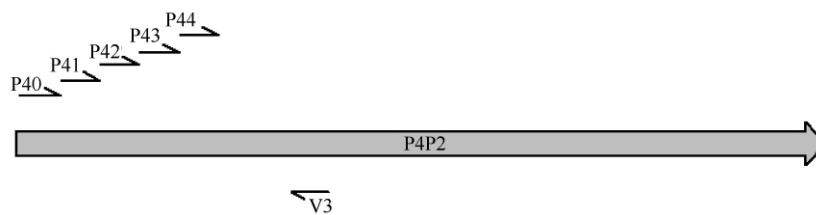


图 1 鉴定 *yigP* 启动子上游边界所用引物位置示意图

Fig. 1 Primers used for 5' truncated promoter fragments.

由图 2 可知, 启动子 5' 端缩短至引物 P43 处时 (即缩短 3 个引物的距离, 每段约 18bp, 总计约 54bp), 其活性仍然较高, 但是接着缩短至引物 P44 处时, 其活性突然大幅下降, 由此推断 *yigP* 启动子的上游边界应该在引物 P43 附近。图中 *adi* 为阳性对照, 来源于大肠杆菌 *adiA* 基因的启动子^[7], 而 *ubiA* 为阴性对照, 来源于 *ubiA* 基因上的一段无关序

1.4 β-半乳糖苷酶测活

β-半乳糖苷酶测活方法以 1995 年 Zhang 等^[8]的方法为基础, 适当调整如下: 待菌体生长至合适阶段 (如未指明, 则 $OD_{600} = 2$), 收集适量菌体, 去上清, 沉淀重新悬浮于 0.1 mol/L 磷酸钠缓冲液中, 超声破碎, $10000 \times g$ 离心 30 min, 收集上清, 取 100 μL 上清, 加入 $100 \times MgCl_2$ 溶液 3 μL, 底物液 66 μL, 0.1 mol/L 磷酸钠缓冲液 121 μL, 于 30°C 保温 30 min 后, 加入 500 μL 1 mol/L Na_2CO_3 终止反应, 分光光度计检测在 420 nm 吸收值。其中底物液为: 60 mmol/L Na_2HPO_4 , 40 mmol/L NaH_2PO_4 , 1 mg/mL ONPG, 2.7 μL/mL β-巯基乙醇。β-半乳糖苷酶酶活按以下公式计算 (Miller): $1000 * \frac{(Abs_{420})}{(cells_inputs [OD_{600}]) * (reaction_time [min])}$

2 结果

2.1 *yigP* 启动子上游边界的确定

我们前期的研究发现, *yigP* 基因的 P4P2 片段具有启动子活性, 因此要对其区域进行一定程度上的确定。为此, 我们将启动子的 3' 端固定在引物 V3 处, 而用一系列相邻的引物 (图 1) 分别扩增获得 5' 端依次缩短的 5 个片段, 将这 5 个片段均克隆至以 *lacZ* 为报告基因的启动子探针质粒 pSP-Z 上, 从而获得 5 个重组质粒 pP40V3-Z, pP41V3-Z, pP42V3-Z, pP43V3-Z, pP44V3-Z, 对含有相应质粒的菌株进行 β-半乳糖苷酶活性分析, 结果见图 2。

另外, 我们还注意到 P41V3, P42V3 的启动子活性比 P43V3 的启动子活性低很多, 这可能是由于在引物 P42 附近存在着某一负调控序列所致, 本文后面将会有进一步的研究。

2.2 *yigP* 启动子-10 区和-35 区的鉴定

对于 *yigP* 基因的启动子区域有了一定程度的确定之后, 然后对其序列进行进一步分析。众所周

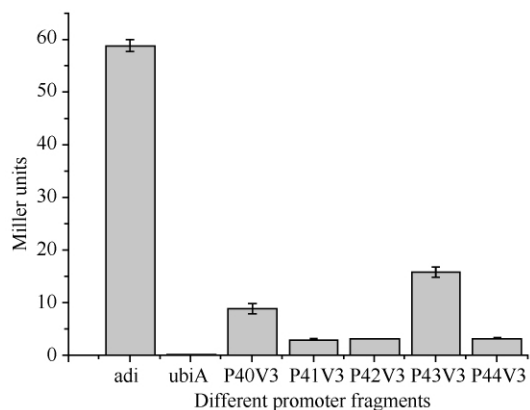


图2 启动子5'端缩短的各片段的活性

Fig. 2 Promoter activity of 5' truncated promoter fragments.

知,启动子最重要的调控序列为-10区和-35区,在原核生物中其保守序列分别为TATAAT和TTGACA。通过对*yigP*启动子的序列进行分析,未发现上述高度保守程度的序列存在。鉴于目前已经用5' RACE技术鉴定了*yigP*的转录起始位点(图3箭头所指,数据未发表),据已了解的启动子序列规律,发现在转录起始位点上游相隔5 bp处有CAGAGT序列,与典型的-10区有3个碱基一致,而在此序列的上游相隔16 bp处又存在着TTCGTT序列,与典型的-35区有2个碱基一致。初步推测上述序列即为*yigP*启动子的-10区和-35区。

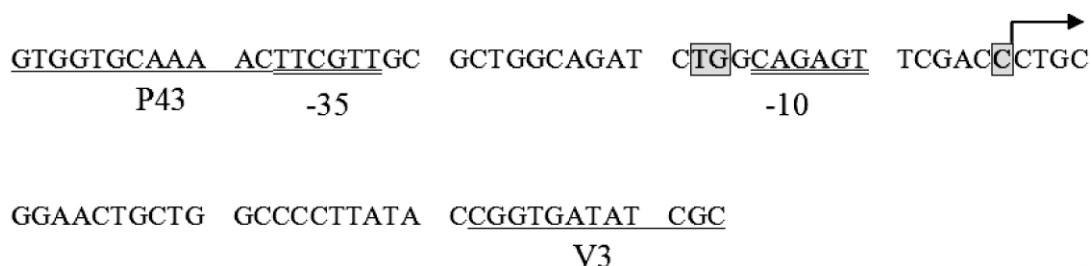


图3 启动子-10区和-35区预测示意图

Fig. 3 Prediction of promoter -10 and -35 region.

为了进一步验证所推测序列的准确性,采用点突变法检测该区域的序列对启动子活性的影响。由于片段P43V3仍然具有相对较高的启动子活性,因此计划在该片段的基础上进行突变实验。

首先,在其它序列均保持不变的情况下,将此可能的-10区序列突变成保守的TATAAT序列和保守程度更低的CAGAGC。将上述含有突变的两个片段分别克隆至启动子探针pSP-Z中,获得两个含有突变的质粒。 β -半乳糖苷酶活性分析显示(图4-A),当该-10区突变成保守的TATAAT时,启动子活性较野生型的增强了2倍,而突变成保守程度较低的CAGAGC时,相应的启动子活性比野生型降低了50%以上。由此可以推断该区域就是*yigP*启动子的-10区。

其次,在其它序列均保持不变的情况下,将此可能的-35区序列突变成保守的TTGACA序列和保守程度更低的GGCGTT。将上述含有突变的两个片段分别克隆至启动子探针pSP-Z中,获得两个含有突变的质粒。 β -半乳糖苷酶活性分析显示(图4-B),当该-35区突变成保守的TTGACA时,启动子活性

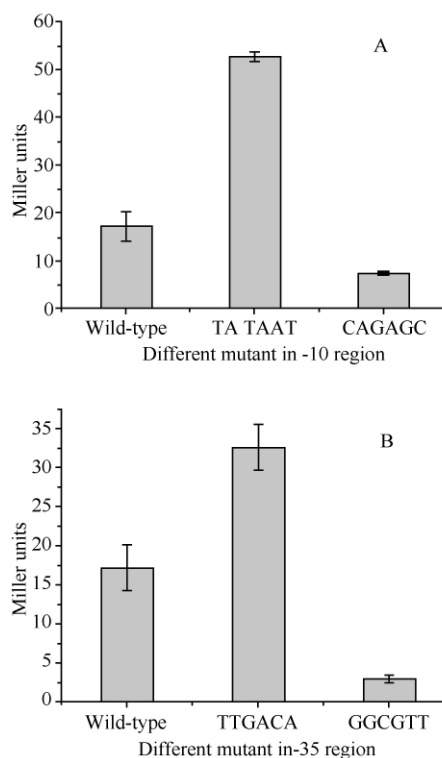


图4 突变-10区或-35区后对启动子活性影响

Fig. 4 Promoter activity effected by mutation in -10 or -35 region.

增强近一倍,而突变成保守程度较低的 GGCGTT 时,相应的启动子活性比野生型大幅下降。由此可以推断该区域就是 *yigP* 启动子的-35 区。

2.3 *yigP* 启动子并非“extended -10”启动子

在大肠杆菌中, σ^{70} 依赖的启动子除了有上述保守的-10 区以外,通常也会含有一个名叫“extended-10”的元件。这种“extended -10”元件通常位于启动子-10 区的上游相隔一个碱基处,其主要的特征序列为 5'-TG-3',该元件对启动子活性的影响很大,任何一个碱基突变都将导致启动子活性的急剧下降。据文献报道,这种启动子在大肠杆菌中约占 20%。通过分析 *yigP* 启动子序列,发现在启动子-10 区上游仅隔一个碱基处有 5'-TG-3'序列(图 3),因此推测 *yigP* 启动子可能为“extended -10”启动子。为了验证这个推测,拟采用点突变方式将此序列由 5'-TG-3'突变成 5'-AC-3',检测其对启动子活性的影响。

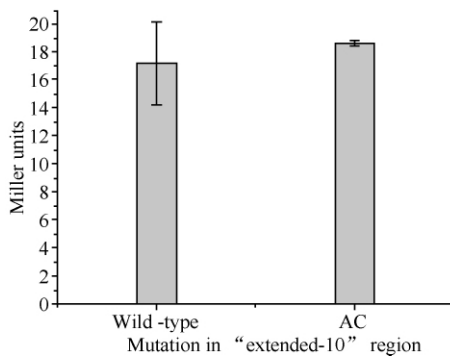


图 5 突变“extended -10”后对启动子活性影响

Fig. 5 Promoter activity effected by mutation in 'extended -10' region.

结果显示(图 5),当“extended -10”区域的 TG 突变成 AC 时,启动子活性无显著变化($P > 0.05$)。即该区域的碱基 TG 对 *yigP* 启动子活性无明显影响,因此可认为 *yigP* 启动子虽然具有“extended-10”启动子的序列特征,但是并非属于“extended-10”启动子。

2.4 P42V3 和 P43V3 片段启动子活性比较

在研究启动子的上游边界时,我们注意到一个比较有意思的现象,即随着启动子 5'端缩短至 P43 区域时,启动子的活性反而增强了(图 2),换句话说,也就是 P42 处序列的存在使得启动子的活性下降。于是,推测引物 P42 处的序列可能是一段负调控序列,对 *yigP* 启动子的活性起着负调控作用。

为了进一步证实片段 P43V3 的启动子活性的确比片段 P42V3 的启动子活性强,先调查了在菌体

生长不同时间段两者的启动子活性强度关系。将菌株 pP43V3-Z/JM83 和 pP42V3-Z/JM83 分别接种过夜,次日以 1% 量翻瓶至 100 mL LB 培养基中,分别培养至 5, 7, 9, 11, 13, 24 h 时,检测其 β -半乳糖苷酶活性。实验结果如图 6。

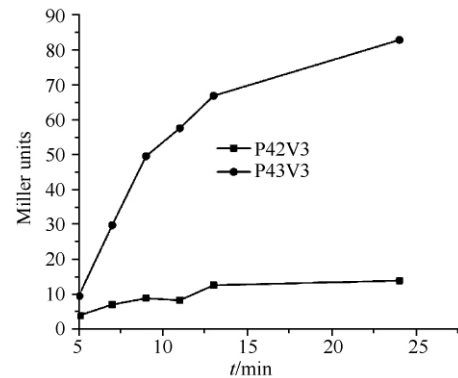


图 6 不同生长时间段 P43V3 和 P42V3 片段启动子活性比较

Fig. 6 Compare of P43V3 and P42V3 promoter activity in different growth phases.

结果显示,在菌体生长的不同时间段,片段 P43V3 的启动子活性都要比片段 P42V3 强,而且随着培养时间的延长,上述两片段间的启动子活性差别越来越大。至此,我们更有理由相信,引物 P42 处的序列是 *yigP* 启动子的负调控序列。

2.5 单碱基突变对该负调控序列的影响

为了继续深入了解这一段负调控序列,我们调查了该序列处单碱基突变对其调控能力的影响,图 7 显示引物 P42 的位置与序列,共 17 个碱基,每个碱基都将其突变成互补的碱基,如 A \rightarrow T, C \rightarrow G。

将上述 17 个菌株以及野生型 pP42V3-Z/JM83, 缺失型 pP43V3-Z/JM83 分别接种过夜,次日以 1% 量翻瓶于 30 mL 培养基中,生长至 $OD = 2$ 时收集菌体,按材料与方法所述检测其 β -半乳糖苷酶活性。各菌株 β -半乳糖苷酶活性见图 3.31。

由图 8 可知,多数位点的碱基点突变后其启动子活性都介于野生型 P42V3(图 8, 0 #)和缺失型 P43V3(图 8, 18#)的启动子活性之间。其中除部分几个位点突变株的启动子活性变化不大外,其余大多数位点的突变都会使启动子活性有所增加,也就是说使得 P42 处序列的负调控能力减弱。此外,也有某些位点,如 7 号 8 号等中间几个位点的突变使得该启动子活性几乎接近于缺失型 P43V3,更是充分说明这些位点对于此负调控序列的重要性,一旦

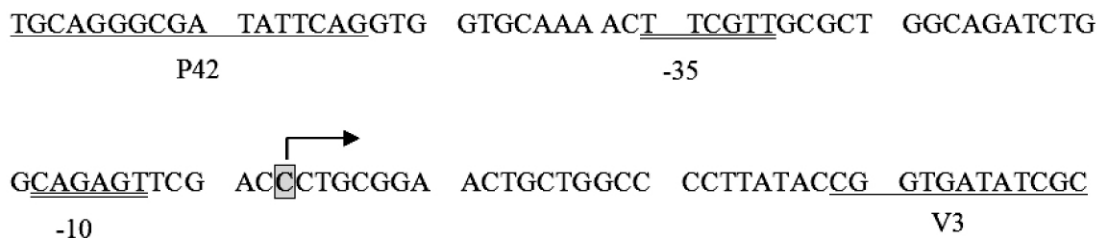


图 7 引物 P42 的位置与序列示意图

Fig. 7 The location and sequence of primer P42.

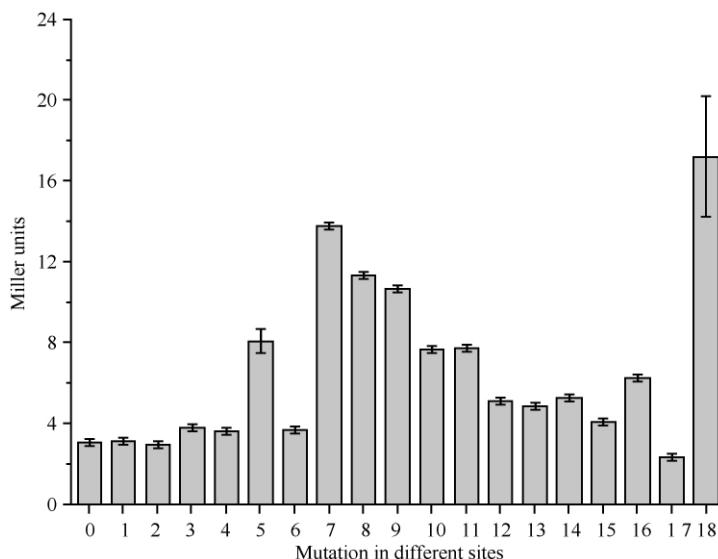


图 8 不同位点点突变对启动子活性的影响

Fig. 8 Promoter activity effected by single base mutation in different sites. 0: P42V3; 1→17: mutations in site 1→17; 18: P43V3.

这些位点发生变化,将导致此负调控序列调控能力的丧失。

3 讨论

yigP 基因的启动子-10区和-35区的保守程度均较低,与典型的启动子序列相比较分别只有3/6、2/6相匹配,因而用普通的启动子分析软件^[9]并未能够预测到。此外,我们检测到的*yigP*启动子活性相对来说比较低,这可能也是由于其保守程度低有关。文献报道的“extended-10”启动子往往是在-10区前间隔一个碱基处具有“TG”序列,而我们的实验证明,并非所有具有这种特征序列的启动子都是“extended-10”启动子,例如*yigP*启动子虽然也具有该特征序列,而该序列对启动子的活性几乎没有影响。

基因的转录一般由启动子介导,同时还可能受到其它元件的调控,或激活,或抑制。例如 lac 操纵

基因^[10],它能在阻遏蛋白的作用下阻遏转录,或在去阻遏物的作用下激活转录。而*yigP*的负调控序列可能与上述操纵基因类似,在某些蛋白的作用下,对基因的表达起负调控作用,对该调控序列进行突变发现位于中间部位的一些碱基对该序列的调控能力影响巨大,可能是其与蛋白相互作用的位点。另外,从图7中可以看到,P42处的负调控序列与启动子的-35区只间隔十几个碱基,一种可能的情况是,存在某一阻遏蛋白与该序列结合,从而影响了RNA聚合酶与启动子的结合,进而影响了基因的转录。然而,对该序列进行分析,并未发现类似于其它操纵基因序列中的对称序列,其具体作用机制有待于进一步的深入研究。

参考文献

- [1] Hawley DK, McClure WR. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids*

- Research*. 1983 ,11:2237-2255.
- [2] Jennie E. Mitchell DZ , Stephen J. W. Busby and Stephen D. Minchin. Identification and analysis of 'extended-10' promoters in *Escherichia coli*. *Nucleic Acids Research*. 2003 ,31:4689-4695.
- [3] Burr T , Mitchell J , Kolb A , Minchin S , Busby S. DNA sequence elements located immediately upstream of the -10 hexamer in *Escherichia coli* promoters: a systematic study. *Nucleic Acids Research*. 2000 ,28:1864-1870.
- [4] Voskuil MI , Voepel K , Chambliss GH. The -16 region , a vital sequence for the utilization of a promoter in *Bacillus subtilis* and *Escherichia coli*. *Molecular Microbiology*. 1995 ,17:271-279.
- [5] Semsey S , Krishna S , Sneppen K , Adhya S. Signal integration in the galactose network of *Escherichia coli*. *Molecular Microbiology*. 2007 ,65:465-476.
- [6] 李月 , 傅楠 , 叶江 , 张惠展. *yigP*——大肠杆菌生长所必需的基因. 华东理工大学学报: 自然科学版 (*Journal of East China University of Science and Technology: Natural Science*). 2011 ,37:453-457.
- [7] 林升阳 , 吴海珍 , 叶江 , 张惠展. 大肠杆菌可诱导启动子表达系统的构建. 上海师范大学学报: 自然科学版 (*Journal of Shanghai Normal University: Natural Sciences*). 2006 ,35:75-79.
- [8] Zhang X , Bremer H. Control of the *Escherichia coli* rrnB P1 promoter strength by ppGpp. *The Journal of Biological Chemistry*. 1995 ,270:11181-11189.
- [9] Munch R , Hiller K , Grote A , Scheer M , Klein J , Schobert M , Jahn D. Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics*. 2005 ,21:4187-4189.
- [10] Gilbert W , Maxam A. The nucleotide sequence of the lac operator. *Proceedings of the National Academy of Sciences*. 1973 ,70:3581.

Identification of transcriptional regulatory sequences of *yigP* gene in *Escherichia coli*

Yi Wang , Jiang Ye , Huizhan Zhang*

State Key Laboratory of Bioreactor Engineering , East China University of Science and Technology , Shanghai 200237 , China

Abstract: [Objective] We investigated the promoter activity of *yigP* gene and analyzed its transcriptional regulatory sequence. [Methods] We cloned promoter fragment into promoter probe plasmid which has a reporter gene *lacZ* , thus the promoter activity could be measured by detecting β -galactosidase activity. Then , the promoter region was minimized by cloning different truncated promoter fragments into promoter probe plasmid. Using site-directed mutagenesis technology , we introduced site mutations in some key sequences and investigated their effects on promoter activity. [Results] The promoter region of *yigP* gene was determined , and also the -10 region and -35 region were identified. Meanwhile , a negative regulatory sequence just upstream of *yigP* promoter was discovered , and our results show that these sequences have important influence on the transcriptional regulation. [Conclusion] We identified the transcriptional regulatory sequences of *yigP* gene , and this facilitated our understanding of gene transcription.

Keywords: *yigP* , transcriptional regulation , promoter , reporter gene , site-directed mutagenesis

(本文责编: 张晓丽)

Supported by the National Natural Science Foundation of China(31070073)

* Corresponding author. Tel: +86-21-64252515; E-mail: huizhzh@ecust.edu.cn

Received: 9 January 2012/Revised: 18 February 2012